

Latent Generativity and Counterfactual Capability: An Identification Framework for Finite Traces

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

A diagnostic for generativity exploitation compares a participant’s effective generative capability across arrangements. Its object is a set of viable trajectories reachable under declared conditions. Nothing observable is that set: what can be recorded is a finite sample of realized trajectories, produced under conditions that actually obtained and through a channel that may project, censor, aggregate, or add noise. Three constraints sharpen the difficulty. A trajectory is an observation of generativity rather than generativity itself, and two distinct dynamical systems can share one locally. Future repertoires cannot be predicted and past generativity is hard to evaluate, because trajectories only sample it. And chained counterfactual reasoning about unrealized generation is a form legal reasoning generally declines. Together these threaten the surrounding programme with unfalsifiability. This paper responds by decomposing the problem into six inferential layers with independent failure modes, and by establishing where evidence does and does not constrain the capability object. The negative half is transferred with attribution from two companion projects, which established that exact equality of accessible history laws fixes binary discrimination at chance, that behaviorally silent structure is invisible, and that enlarging a probe family refines an equivalence partition. The paper’s own contribution concerns a target those results do not reach: a reachable repertoire is set-valued and is indexed to conditions that were never realized. Writing the repertoire as a superlevel set of a reachability functional converts a question about set equality into a question about how far that functional moves across an observational equivalence class. Five results follow. The counterfactual repertoire is sharply bounded by the core and hull of the class. A model class permitting on-regime behavior to be spliced to arbitrary off-regime behavior collapses those bounds to whatever the class alone permits, so the lift fails maximally. Otherwise uncertainty about the organization confines uncertainty about the repertoire to a band around the reachability threshold, whose width is the class’s off-regime deviation. A capability comparison is identified across the entire class whenever the gap between assessed and baseline functionals exceeds the combined ambiguity. And under uniformly condition-Lipschitz kernels that ambiguity is at most twice the Lipschitz constant times the distance from the counterfactual condition to the observed regime. The surgery construction and the Lipschitz bound bracket the problem: without a structural tie between off-regime and on-regime behavior no counterfactual information is available, and with a uniform tie the bounds degrade gracefully with counterfactual distance. Four theses argue that non-identifiability is a property of a model class paired with a design rather than evidence against the object, that the framework is falsifiable exactly because it returns determinate negatives, that identification failure allocates an evidential benefit to the party controlling the records, and that existence claims reach conclusions magnitude claims cannot. The paper establishes no estimator, no general non-identifiability theorem, and no valuation method, and classifies no observed relation.

Keywords: identification; partial identification; observational equivalence; counterfactual capability; reachable repertoire; generativity exploitation; evidential burden.

Document Status, Preparation, and Reuse

This unnumbered section records the manuscript’s status, research genealogy, preparation process, and reuse conditions. These statements govern the document and carry no evidential weight for its substantive position.

Status and correspondence. This paper is a revisable discussion document. Its definitions are proposals, its mathematical results are conditional on declared assumptions, and its empirical bridges remain open. Objections, counterexamples, corrections, alternative formulations, and relevant literature are welcome at huangwanhong@serendip.ngo.

Research genealogy and transfer. The author-led exploratory discussion supplied three constraints: that a trajectory is an observation of generativity rather than generativity itself and that two dynamical systems can share one locally; that future repertoires cannot be predicted and past generativity is hard to evaluate; and that reasoning from a hypothetical chain of unrealized generation to an unrealized sum is a form legal reasoning declines. These constraints generate the paper and are not premises of any result.

Section 7 transfers results established for two companion projects in the same research programme rather than proving them anew. The exact-equivalence discrimination bound and its total-variation refinement were proved for a brain-in-a-vat experiment; the silent-factor and arcsine constructions and the probe-family monotonicity result were proved for a study of topological warrant in developmental language. The transfer is stated explicitly at each point of use. Section 8 contains the results original to this paper.

Generative AI use. Generative AI systems were used in the preparation of this work. Anthropic’s Claude and OpenAI systems including ChatGPT and Codex supported exploratory discussion, formal reconstruction, source discovery followed by verification, and drafting in $\text{\LaTeX}$. The author determined the research questions, theoretical commitments, and epistemic status of the claims and bears sole responsibility for the manuscript, including its errors. None of these systems is an author or holds authorship credit.

License. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). The license text is available at <https://creativecommons.org/licenses/by-nc/4.0/>.

1 Introduction

A diagnostic that compares a participant’s effective generative capability across arrangements has an awkward object. Capability, on the construction it uses, is a set of viable trajectories reachable from the participant’s actual position under a declared arrangement. Nothing observable is that set. What can be recorded is a finite sample of realized trajectories, produced under conditions that actually obtained, through a channel that may project, censor, aggregate, or add noise. The object is a set of possibilities under conditions that did not obtain; the evidence is a sample of actualities under conditions that did.

Three constraints sharpen the mismatch. A trajectory is an observation of generativity rather than generativity itself, and two distinct dynamical systems can produce the same trajectory over a limited region, so resemblance of outputs supports no claim about shared organization. Future repertoires cannot be predicted and past generativity is hard to evaluate, because trajectories only sample it. And reasoning of the form *had this participant not been affected, it would have generated one thing, then another, then earned a given sum* is a form legal reasoning generally declines, since such chains compound unverifiable steps.

Together these threaten the surrounding research programme with unfalsifiability. A diagnostic whose central object is never accessible can return only positive or unresolved verdicts, and a test that cannot fail is not a test. That threat, rather than any application, is what this paper addresses.

The response is neither to solve the identification problem nor to concede it, but to decompose it and then establish, layer by layer, where evidence does and does not constrain the object. Six inferential layers are separated: trace evidence, latent-class identification, reachable-repertoire identification, feasible-baseline comparison, benefit-through identification, and magnitude. Each fails independently, and a reporting rule prohibits three specific promotions across them, since provenance does not establish identity, a repertoire contraction does not establish a benefit path, and an existence finding does not establish a valuation.

Two of the paper’s results are transferred rather than new, and the boundary is stated explicitly wherever they are used. Companion projects in the same programme have established that exact equality of accessible history laws under every admissible policy fixes binary discrimination at chance, with a total-variation ceiling for approximately matched cases; that behaviorally silent structure is invisible, through an explicit augmentation construction and a one-variable instance; and that enlarging a probe family refines an equivalence partition. Adapted here with a design family in place of a policy class, these establish both directions of the first constraint: trace similarity licenses no identity claim, and trace dissimilarity licenses no independence claim, since dissimilarity confounds a difference in organization with a difference in conditions.

The paper’s own contribution concerns a target those results do not reach. They answer questions about which model generated observed data. A reachable repertoire is a *set* of trajectories under conditions that were *never realized*, and non-identification of the generating model is neither necessary nor sufficient for non-identification of that set. Writing the repertoire as a superlevel set of a reachability functional converts a question about equality of sets into a question about how far a scalar functional can move across an observational equivalence class, and five results follow. The counterfactual repertoire is sharply bounded by the core and hull of the class. A model class permitting on-regime behavior to be spliced to arbitrary off-regime behavior collapses those bounds to whatever the class alone permits, so on-regime evidence constrains the counterfactual exactly as much as no evidence. Otherwise uncertainty about the organization confines uncertainty about the repertoire to a band around the reachability threshold, so membership is entailed outside that band despite the organization being unidentified. A capability comparison is identified across the entire class whenever the gap between assessed and baseline reachability exceeds the combined

ambiguity. And under uniformly condition-Lipschitz kernels the ambiguity is at most twice the Lipschitz constant times the distance from the counterfactual condition to the observed regime. The surgery construction and the Lipschitz bound bracket the problem: without a structural tie between off-regime and on-regime behavior no counterfactual information is available, and with a uniform tie the bounds degrade gracefully with counterfactual distance.

Four further insights are argued rather than derived. Non-identifiability is a property of a model class paired with a design, so it is evidence about evidence rather than about the object. The framework is falsifiable exactly because it returns determinate negatives, and a design’s discriminating power is itself a reportable quantity, so an underpowered design can be shown to be underpowered. Identification failure is not evidentially neutral between the parties, since the most measurable layer is the one whose records are held by the party setting the terms, and the affected participant’s own evidence is precisely the evidence that licenses no inference. And a supported existence claim reaches conclusions an unsupported magnitude claim cannot, which locates what survives the refusal of chained counterfactual valuation.

The paper establishes no estimator, no general non-identifiability theorem, and no valuation method, and it classifies no observed relation. Its results are identification results rather than inference procedures, and they rest on idealizations that realistic model classes may not satisfy.

The argument proceeds in ten further sections. Section 2 states the scope and the three constraints. Section 3 performs the problem analysis and states the layer decomposition. Section 4 reviews traditions that confront underdetermined targets. Section 5 argues the four theses. Section 6 develops the political economy of evidence. Section 7 transfers the non-identification results. Section 8 proves the paper’s own results. Section 9 supplies a complementary bound that bypasses latent identification. Section 10 specifies a worked design, Section 11 collects open questions, and Section 12 states limits and defeat conditions.

The philosophical and formal parts are *formally independent*, since every proposition follows from declared assumptions and no derivation cites a thesis, and *justificatorily dependent*, since the formal results are a schema whose claim to be worth constructing rests on the theses. Section 5.5 records what fails if each is rejected.

2 Inferential Scope and Target Separation

This section states the problem, the constraints that generate it, the threat they pose to the surrounding programme, and the paper’s stopping point. It closes by fixing the relation between the paper’s philosophical and formal parts.

2.1 The Object and What Can Be Observed

This subsection states the mismatch the paper addresses.

A companion diagnostic evaluates whether a participant’s effective generative capability is materially worse under an assessed arrangement than under an independently admissible comparison. Its primary object is a reachable viable repertoire: the set of trajectories the participant could realize, from its actual position, under a declared arrangement, horizon, and viability standard.

Nothing observable is that set. Observation supplies a finite sample of realized trajectories, produced under conditions that actually obtained, and recorded through a channel that may project onto a lower-dimensional summary, censor whole regions, aggregate across episodes, or add noise. The diagnostic’s object is a set of possibilities under conditions that did not obtain. The evidence is a sample of actualities under conditions that did.

2.2 Three Constraints

This subsection states the constraints that generate the paper. Each was raised in the source discussion as an objection to a formalization then under construction, and each rules something out.

A trajectory is not a generativity. Two distinct dynamical systems can produce the same trajectory over a limited region. A rule inferring shared generative organization from resemblance of outputs, reasoning steps, or workflows would therefore be unsafe. This rules out any construction treating an observed trace as the capability object. It was raised as a representational problem: an instrument that fixes on the trajectory has not fixed on the object.

Evaluation fails in both temporal directions. Future generative repertoires cannot be predicted, and past generativity is hard to evaluate, because trajectories only sample or approximate it. This separates two problems frequently merged: prospective uncertainty about what would have been produced, and retrospective non-identifiability of what was there. It rules out treating the difficulty as solvable by better record-keeping alone.

Counterfactual chains are not accepted forms of reasoning. An argument of the form *had the participant not been affected, it would have generated this, then that, then earned a given sum* is a form legal reasoning generally declines, and the reluctance is not arbitrary: such chains compound unverifiable steps. This rules out founding the framework’s practical relevance on chained counterfactual valuation.

2.3 The Threat and the Response

This subsection states what the constraints endanger and how the paper responds.

Together the constraints threaten the surrounding programme in a specific way. If the diagnostic’s central object is never accessible, the diagnostic can return only positive or unresolved verdicts, and a test that cannot fail is not a test. The programme would then rest on a concept that no evidence could discipline.

The response is neither to solve the identification problem nor to concede it. It is to decompose it. Section 3 separates six inferential layers with independent failure modes and states a reporting rule prohibiting promotion across them. Sections 7 through 10 then establish, layer by layer, where evidence constrains the object and where it does not. The central positive finding is that latent non-identification does not entail capability non-identification, and that the conditions separating the two cases are statable.

Claim 2.1 (Limited inferential position). Relative to a declared model class, observation-design family, horizon, and condition metric, a capability comparison across arrangements can be identified without identifying the generative organization, exactly when the difference between assessed and baseline reachability exceeds the residual ambiguity of the organization’s observational equivalence class.

Claim 2.1 is a conditional statement about a model, not a claim that any actual capability has been identified. Its ingredients carry modeling commitments that Section 12 states as defeat conditions.

2.4 Source Use and the Transfer Boundary

This subsection states two constraints governing the material below.

Where a source is used, the text states the specific constraint it supplies and the inference it withholds. The verified corpus establishes formal analogues, methodological constraints, and one legal analogue; it establishes neither the novelty nor the empirical adequacy of the constructions proposed here, and no source is cited in support of a proposition.

Where a result is transferred from a companion project, the transfer is stated at the point of use and the assumptions it carries are named. This matters because the negative half of the paper is largely inherited, and a presentation that blurred the boundary would claim as new what the surrounding programme already established.

3 The Structure of the Identification Problem

This section performs the problem analysis. Its objective is to separate inferential targets that fail independently, and to state the reporting rule that keeps them separate.

3.1 Two Temporal Directions

This subsection separates the problems the second constraint merges.

Ex ante unpredictability concerns what a generative organization would produce going forward. It is a problem about the future, and it persists even for an organization that is perfectly known: a fully specified stochastic process still has an unrealized future.

Ex post non-identifiability concerns what organization was there. It is a problem about inference from a record to its source, and it persists even where the future is irrelevant, as when assessing a past foreclosure.

The two have different remedies and different failure modes. Richer observation addresses the second and leaves the first untouched. A longer horizon addresses neither. Conflating them produces the impression that the identification problem is a single obstacle, when in fact only one of the two is an inference problem at all.

3.2 Six Inferential Layers

This subsection states the paper’s central methodological proposal.

Table 1 records six layers, each with its own evidence requirement and its own failure mode. The proposal is that a report must supply a status for each layer, and that success at one layer promotes nothing at another.

3.3 Three Prohibited Promotions

This subsection states the reporting rule by naming the errors it blocks.

Provenance to identity. Evidence that one participant had access to another’s work supports a historical claim about access. It supports no claim that the two share a generative organization. Section 7 makes this precise.

Table 1: Inferential Layers, Their Evidence Requirements, and Their Failure Modes. The table organizes obligations and supplies no evidence. Layer coupling is evaluative rather than logical: later layers take earlier outputs as inputs, while success at one layer entails success at none.

Layer	Evidence requirement	Characteristic failure
Trace evidence	Records of realized trajectories, with conditions, channel, resolution, and horizon declared	Channel unrecorded, so the projection cannot be inverted or bounded
Latent-class identification	A design family rich enough to separate candidate organizations	Equivalence class is not a singleton; distinct organizations remain indistinguishable
Reachable-repertoire identification	Behavior of the equivalence class under conditions not realized	Class members agree on-regime and diverge off it
Feasible-baseline comparison	A technically feasible and independently admissible alternative arrangement	No admissible comparison, or comparison selected by its own conclusion
Benefit-through identification	A structural model with intervention, mediator, and identification assumptions	Coincident loss and gain with no path between them
Magnitude and valuation	A valuation functional over foreclosed trajectories	Existence supported while worth remains unidentified

Repertoire contraction to benefit-through. Evidence that a participant’s reachable repertoire contracted supports no claim that another party’s advantage runs through that contraction. Causal mediation methodology separates the definition of such an effect from its identification, estimation, and sensitivity analysis [6], and the separation applies here: the contraction is one layer and the path is another.

Existence to valuation. Evidence that a contraction occurred supports no claim about what the foreclosed trajectories were worth. Section 5 argues that this separation has normative content rather than being a mere caution.

The reporting rule requires that an unresolved layer be recorded as unresolved rather than as a negative finding. This is the operational form of the three-valued discipline the companion diagnostic uses.

3.4 Requirements on a Replacement Account

This subsection converts the analysis into obligations discharged below.

The account needs a precise statement of what observation leaves open at the second layer; a statement of when that openness does and does not propagate to the third; a route to capability claims that bypasses the second layer entirely, for use when it fails; and a design specification stating what would have to be measured. Sections 7 through 10 supply these in order.

4 Identification Traditions and Rival Approaches

This section locates the paper among traditions that confront underdetermined targets. Its objective is bounded positioning: each supplies a formal analogue or a methodological constraint, none supplies a result about generative capability, and the organizing finding concerns how they respond to underdetermination.

4.1 Partial Identification

This subsection identifies the tradition whose posture the paper adopts.

Manski derives worst-case nonparametric bounds on treatment effects, establishing that informative statements remain available where point identification fails, and that the resulting bounds are sharp given the stated assumptions [10]. Tamer surveys the resulting research programme, including the systematic trade-off between assumption strength and interval width and the practice of reporting identified sets in place of point estimates [15].

Molchanov and Molinari develop random-set theory as a framework for identification analysis when data and assumptions reveal a set containing the parameter rather than its value [11]. This is the closest available vocabulary for the present target, since a reachable repertoire is set-valued by construction. The paper uses the vocabulary of core and hull and asserts no random-set theorem.

Two limits apply. The econometric setting has a well-defined parameter and a sampling process; the present setting has a set-valued object under a condition regime that was never realized, which is a harder target. And a standard criticism of worst-case bounds is that they are frequently too wide to be useful. Section 12 records that this paper inherits the risk.

4.2 Structural Identifiability and Observational Equivalence

This subsection identifies the formal analogues for the equivalence class.

Bellman and Åström analyze structural identifiability: whether the parameters of a model structure can in principle be recovered from ideal input–output data, independently of noise or sample size [1]. The distinction between structural identifiability, a property of the model structure paired with an experiment, and practical identifiability, a property of finite noisy data, is the distinction this paper requires when it indexes its equivalence class to a design family.

Givan, Dean, and Greig develop equivalence notions and model minimization for Markov decision processes, characterizing when distinct models are indistinguishable with respect to a class of behaviors [5]. This supplies the standard a well-posed equivalence notion should meet. Kaelbling, Littman, and Cassandra formalize acting when the state is observed only through a partial and noisy channel [7], which is the structure of the observation channel used below.

None of these addresses generative organizations, capability, or counterfactual repertoires, and no result from them transfers. What transfers is the framing.

4.3 Underdetermination Proved Rather Than Assumed

This subsection identifies the methodological posture the paper adopts toward its own negative results.

Shalizi and Thomas prove that homophily and contagion are generically confounded in observational network studies [14]. The paper takes this as a model of the correct response to an underdetermined target: prove the non-identification rather than weakening an estimate and proceeding.

4.4 Capability and the Limits of Observed Choice

This subsection identifies why observed behavior underdetermines the object.

Sen distinguishes achieved functionings from the capability set of alternatives a person could feasibly achieve and choose [13]. The identification problem follows directly: what is observed is achievement, and what the diagnostic requires is the set. Robeyns records the choices required to operationalize any capability analysis, including dimension selection, weighting, and the admissibility

Table 2: Traditions Confronting Underdetermined Targets. Every tradition reviewed responds to underdetermination by weakening the claim rather than by strengthening the inference. This paper adopts the same posture.

Tradition	Underdetermined target	Constraint imposed here
Partial identification [10, 15]	Treatment effect under weak assumptions	Report identified sets; accept that weak assumptions may yield uninformative width
Random sets [11]	A set containing the parameter	Supplies core and hull vocabulary; no theorem is inherited
Structural identifiability [1]	Model parameters from ideal data	Non-identifiability is a property of structure paired with experiment, not of the world
Equivalence notions [5]	Models indistinguishable for a behavior class	A well-posed equivalence notion must be relative to a declared class
Partial observability [7]	State behind a channel	The channel must be declared before the projection is inverted or bounded
Confounding results [14]	Influence against homophily	Prove the non-identification; do not weaken an estimate and proceed
Capability measurement [13, 12]	Opportunity behind achievement	Declare dimensions, weights, and incompleteness; report their variation
Adaptive preferences [4]	Feasibility behind behavior	Behavioral inference of feasibility is inadmissible under sustained constraint
Mediation discipline [6]	Effect through an intermediate	Definition, identification, estimation, and sensitivity remain separate

of incompleteness [12], which is why the specification below is a declared input whose variation must be reported.

Elster analyzes the adjustment of preferences to perceived feasibility, whereby an agent ceases to want what it has come to regard as unavailable [4]. This matters acutely here. Where feasibility is inferred from behavior, preference adaptation corrupts the measurement, and it does so worst in exactly the cases the diagnostic targets, since sustained constraint is what produces adaptation. Section 9 accordingly restricts its result to directly documented condition changes.

4.5 Comparative Summary

Table 2 records each tradition’s underdetermined target, its response, and the constraint it imposes here.

5 Epistemic and Philosophical Foundations

This section argues the four theses on which the paper’s significance rests. Each subsection states a thesis, argues for it, states the strongest objection known to the author, and replies. The theses answer questions the formalism cannot answer for itself: what a non-identification result shows about the object, whether the surrounding framework is falsifiable, whether identification failure is evidentially neutral between parties, and what an existence claim without a magnitude claim can support.

5.1 Non-Identifiability as an Epistemic Relation

This subsection blocks the strongest misreading of the paper’s negative results.

Thesis 5.1 (Epistemic rather than ontological limitation). Non-identifiability is a property of the relation between a model class and an observation design. It is not evidence that the latent organization is unreal, that capability talk is empty, or that the diagnostic’s object was badly chosen.

The argument is that every non-identification statement below is doubly indexed. The equivalence class of Definition 7.1 is a function of the model class and the design family. Enriching the design family refines the class, by a monotonicity result proved for a companion project and stated as Proposition 7.6; restricting the model class refines it equally. Neither operation changes anything about the participant. An inference from “indistinguishable under this design” to “not there” would license the same conclusion about any quantity whose current measurement is coarse.

The strongest objection is that a disposition no design could ever distinguish is idle, and that postulating it violates parsimony. The reply concedes the principle and denies its application. Section 8 exhibits a class of claims about the object that observation does constrain: outside an explicit band, membership in the counterfactual repertoire is entailed by the evidence. An idle posit would admit no such claims. The object earns its place by supporting bounded inference, not by assertion.

5.2 Falsifiability Through Determinate Negatives

This subsection answers the objection that would otherwise defeat the surrounding programme.

Thesis 5.2 (Answerable unfalsifiability). A framework that frequently returns unresolved verdicts is falsifiable exactly when it can also return determinate negatives. The results below supply that capacity in a specific and quantitative form.

The objection to be answered is that a diagnostic yielding only positive and unresolved verdicts is unfalsifiable and therefore empty. The reply is that the framework returns determinate negatives in stated circumstances. The false-negative direction of Section 7 establishes that an apparent contraction can be shown to reflect a condition difference rather than a capability difference. Corollary 8.7 has a negative counterpart: when the assessed reachability functional dominates the baseline functional by more than the combined ambiguity, no member of the equivalence class exhibits a disadvantage, and the framework returns a supported negative.

The transferred bound of Proposition 7.2 strengthens this. Residual distinguishability between candidates sets an explicit ceiling on any discrimination procedure, so a design’s discriminating power is a reportable quantity. A framework that reports its own ceiling can be shown to be underpowered, which is a way of being wrong.

The strongest objection is that determinate negatives may be practically rare, so the capacity is formal rather than real. The reply concedes this and converts it into a commitment: Section 12 records that if no negative case is ever produced in a domain where the framework is applied, that outcome counts against the framework rather than being absorbed as caution.

5.3 The Evidential Asymmetry

This subsection argues the paper’s most contestable claim and supplies its political-economic content.

Thesis 5.3 (Asymmetric epistemic position). Identification failure is not evidentially neutral between the parties. The actor controlling realization conditions typically also controls the records that would identify a contraction, so the default consequence of non-identifiability systematically favors that actor.

The argument proceeds from the structure of the results below. Section 9 establishes that the most tractable route to a capability claim runs through directly documented contraction of the feasible action set. Section 4 establishes that the alternative, inferring feasibility from behavior, is corrupted by preference adaptation precisely under sustained constraint. So the usable route requires documentation of access terms, permission changes, allocation decisions, and pricing. Those records are held by whoever sets them.

The affected participant’s own evidence is largely its realized trajectory, which Section 7 shows licenses no inference about the organization. The layer that is most identifiable is therefore the layer whose records are least accessible to the party with the complaint. A regime resolving non-identifiability against the complaining party allocates the benefit of an evidential gap to the party that produced it.

The strongest objection is that this is a general argument for discovery obligations with nothing specific to generativity, and that it proves too much. The reply concedes the general point and identifies what survives it. What is distinctive is that the object of the claim is constitutively unobservable to the complainant. Reachable repertoires under alternative conditions are not the kind of thing a participant observes about itself: the participant observes what it did, not what it could have done under conditions that did not obtain. The evidence is not merely held by the other side; it was never in anyone’s possession, and the nearest available substitute is documentation the other side holds.

A second objection is that a burden-shifting rule would penalize a party for the intrinsic difficulty of a claim, converting non-identifiability into a presumption of liability. The reply restricts the thesis. It concerns justificatory burden in the sense the companion diagnostic uses, not legal proof. Doctrinal implications are outside this paper. The thesis must also be paired with the constraint against inferring a classification from structural position [17]: an evidential asymmetry never substitutes for the mechanism and benefit conditions of the diagnostic, which a domination-based account would locate in the relation itself [16].

5.4 What an Existence Claim Supports

This subsection determines which conclusions the paper’s evidence can reach.

Thesis 5.4 (Existence without magnitude). A supported existence claim can ground conclusions that an unsupported magnitude claim cannot. Conflating the two produces overreach in one direction and unwarranted dismissal in the other.

The formal argument is that the two come apart in the results below. Corollary 8.7 supports an inclusion between repertoires without identifying either, and Proposition 9.2 supports a comparison without valuing any foreclosed trajectory. Neither yields a magnitude.

The normative argument is that they support different conclusions. An existence finding can support a demand for disclosure, a prospective obligation to alter a mechanism, or a shift in justificatory burden, none of which requires knowing what the foreclosed trajectories were worth. Compensation is the conclusion that genuinely requires magnitude. The third constraint of Section 2 observed that chained counterfactual valuation is a form legal reasoning declines; this thesis locates what survives that refusal.

The legal analogue is instructive and must be handled carefully. King argues that where conduct destroys a chance of a favorable outcome, the appropriate object of the claim is the lost chance itself rather than the outcome, so that a claim can proceed where proof of the outcome on the balance of probabilities is unavailable [8]. The structural point is that legal reasoning has confronted an interest in unrealized futures and has at least once reconceived the object of the claim to preserve

provability. That is the analogy, and it is the whole of it. The doctrine concerns injury with a defined adverse outcome and an estimable probability of avoiding it; generative capability has neither. This paper characterizes no jurisdiction’s law, proposes no doctrine, and does not assert that lost-chance reasoning extends to generativity.

The strongest objection is that without magnitude, trivial and serious contractions are indistinguishable, so existence alone supports no obligation. The reply is that materiality operates on the existence claim directly, through declared resolution on the capability coordinates, and does not require valuing foreclosed trajectories. The distinction between trivial and serious is available before the distinction between worth- x and worth- y .

5.5 Dependence of the Framework on the Four Theses

This subsection records what fails if each thesis is rejected.

Rejecting Thesis 5.1 permits the negative results to be read as eliminating the diagnostic’s object, which would end the surrounding programme rather than discipline it. Rejecting Thesis 5.2 leaves the unfalsifiability objection unanswered, which is the strongest available objection to the programme as a whole. Rejecting Thesis 5.3 removes the paper’s normative and political-economic content, leaving a purely methodological contribution that remains defensible at reduced scope. Rejecting Thesis 5.4 collapses existence into magnitude, which either inflates what the evidence supports or discards evidence that supports something.

Theses 5.2 and 5.3 are load-bearing. Thesis 5.3 is the most vulnerable, and the author would welcome direct challenge to it. Every proposition below survives the rejection of all four.

6 Opacity, Records, and the Political Economy of Evidence

This section develops the consequences of Thesis 5.3. Its objective is to describe the structural position of the records the framework requires, and to state the practical problem that position creates.

6.1 Where the Usable Records Sit

This subsection locates the evidence the framework needs.

Birch analyzes contemporary technoscientific capitalism as organized around rentiership: the appropriation of value through ownership and control rights, monopoly conditions, and market or regulatory devices [2]. Control rights are exactly the devices whose exercise produces the records Section 9 requires. An access term, a permission change, an allocation decision, and a price are all records of a control right being exercised, and they are generated and retained by the party holding it.

This is not incidental to the domain. The framework’s most tractable design requires documentation of condition changes; condition changes are what control rights effect; and the controller is the party that documents them. The correlation between analytic tractability and inaccessibility is structural.

6.2 Two Kinds of Opacity

This subsection distinguishes cases that call for different responses.

Incidental opacity arises when records exist but are not shared, or when measurement was never undertaken. It is remediable by disclosure requirements, retention rules, or independent measurement.

Constitutive opacity arises when the object was never recorded by anyone, because it is not the kind of thing a participant observes about itself. A participant’s reachable repertoire under conditions that did not obtain has no record anywhere. No disclosure regime produces it.

The distinction matters for what disclosure achieves. Documented condition changes address the third layer of Table 1 and leave latent-class identification, benefit-through, and magnitude untouched. A disclosure regime would therefore be genuinely useful and sharply limited, and a proposal presenting it as a solution to the identification problem would overstate its reach.

6.3 The Practical Problem

This subsection states the difficulty the formal results do not resolve.

The framework’s most usable route to a capability claim is least accessible to the party that would invoke it. A complainant possesses its own realized trajectory, which licenses no inference about the organization, and lacks the documentation that would support the bypass route. This is a practical problem about the framework’s applicability, and no result below addresses it.

The paper records it rather than solving it. Burden allocation in law, the design of disclosure obligations, and the institutional question of who should be required to retain what all belong to a separate project with its own verified sources. What this paper establishes is that the problem is structural rather than a contingent feature of current record-keeping practice.

7 Observational Equivalence and the Failure of Both Inferential Directions

This section establishes what finite traces cannot show. Its results are transferred with adaptation from two companion projects in the same research programme, and the transfer is stated at each point. Its objective is to make precise the first constraint of Section 2.

7.1 Objects and Declarations

This subsection fixes the objects used through Section 9.

Fix a domain d , a model class $\mathfrak{G}_d$ of candidate generative organizations, a carrier, a history h , a horizon H , a reachability threshold $\epsilon \in (0, 1]$, and a resolution $\delta > 0$. Let $(\mathcal{T}, \mathcal{F})$ be a measurable trajectory space whose σ -algebra contains the closed neighborhoods $\mathcal{N}_\delta(\tau)$. Let $\mathcal{C}$ be a space of realization conditions with metric ρ_C , and let $\mathcal{C}_{\text{obs}} \subseteq \mathcal{C}$ be the conditions under which observation occurred.

For each $G \in \mathfrak{G}_d$, condition C , and action a , let $K_G(\cdot \mid h, a, C)$ be a probability kernel on $\mathcal{T}$. Let $\mathcal{A}(C, h)$ be a nonempty feasible action set and $\mathcal{V}(C, h) \in \mathcal{F}$ a viability event; both depend on conditions and history alone, not on G . Let $\mathcal{Q}_0$ be a declared observation-design family, with channel O_q producing finite observations Z from a trajectory for each $q \in \mathcal{Q}_0$.

Definition 7.1 (Observational equivalence class).

$$[G]_{\mathcal{Q}_0, \mathcal{C}_{\text{obs}}} = \{G' \in \mathfrak{G}_d : \mathcal{L}(Z \mid G', C, q) = \mathcal{L}(Z \mid G, C, q) \ \forall q \in \mathcal{Q}_0, \ \forall C \in \mathcal{C}_{\text{obs}}\}.$$

The class is doubly indexed. Nothing in it is a property of the participant alone, which is the formal content of Thesis 5.1.

7.2 Trace Similarity Licenses No Identity Claim

This subsection states the first direction, transferred.

Proposition 7.2 (Exact-equivalence discrimination bound). *Let θ, θ' induce identical accessible history laws under every admissible policy in a declared class Π through horizon T . Then under equal binary priors, every decision rule measurable in the accessible history succeeds with probability exactly $\frac{1}{2}$. Under approximate matching, the optimal success probability is*

$$P_{\text{succ}}^* = \frac{1}{2} \left(1 + \sup_{\pi \in \Pi} \text{TV}(\mathbb{P}_{\theta, \pi}^T, \mathbb{P}_{\theta', \pi}^T) \right).$$

This result was proved for a companion brain-in-a-vat project and is used here without reproof. Its adaptation replaces the policy class Π with the design family $\mathcal{Q}_0$: a design is a way of interrogating the system, exactly as a policy is. The adaptation carries the original assumptions, and two of them bind. The bound is stated for binary hypotheses with equal priors, so its extension to a model class of more than two candidates requires a separate argument. And it presumes a fixed interface, so a design family that changes what is recordable requires re-derivation.

Two constructions instantiate the failure concretely, both proved for a companion project on topological warrant.

Proposition 7.3 (Silent-factor augmentation). *Let a latent process have state x_t , kernel K , and output $y_t = h(x_t)$. Adjoin a coordinate z_t evolving under an arbitrary kernel independently of its own past, and define the output map to ignore it. Then every observable finite-dimensional distribution of the augmented process agrees with that of the original, while the latent state space differs.*

Proposition 7.4 (Arcsine pair). *Let X have the arcsine density on $[-1, 1]$ and let $Y = \cos \Theta$ with Θ uniform on $[0, 2\pi)$. Then X and Y are identically distributed while their supports differ topologically.*

Proposition 7.3 is a precise formalization of the first constraint of Section 2: two different dynamical systems share every observable trajectory law. Proposition 7.4 gives a one-variable instance requiring no process notation. Together they establish that trace similarity, however extensive, supports no claim that two participants share a generative organization. A rule inferring copying from resemblance of outputs would be unsafe in exactly the way the source discussion anticipated.

7.3 Trace Dissimilarity Licenses No Independence Claim

This subsection states the converse direction.

Proposition 7.5 (Condition confounding). *For a single G , there exist conditions and designs with*

$$\mathcal{L}(Z \mid G, C_1, q_1) \neq \mathcal{L}(Z \mid G, C_2, q_2).$$

Proof. Immediate from the dependence of K_G on C and of the channel on q . Any G whose kernel is non-constant in C , together with any two conditions at which it differs, supplies an instance. $\square$

The proposition is elementary and its consequence is not. Observed dissimilarity of expressions confounds a difference in generative organization with a difference in history, tools, permissions, collaborators, or institutional position. Two participants producing visibly different work may share an organization and differ in conditions. This blocks any inference from dissimilarity to independence, and it is the reason the diagnostic must compare arrangements rather than participants.

7.4 Design Enrichment Refines the Class

This subsection states the monotonicity result that gives Thesis 5.1 formal content.

Proposition 7.6 (Probe-family monotonicity). *For admissible design families $\mathcal{Q}_1 \subseteq \mathcal{Q}_2$ using the same response metrics on shared conditions, the induced exact-equivalence partitions satisfy $\Pi_{\mathcal{Q}_2} \preceq \Pi_{\mathcal{Q}_1}$: the partition induced by the larger family refines that induced by the smaller.*

This result was proved for the companion topological project and is used here without reproof. Its consequence is that non-identifiability is never a terminal fact. A richer design family can only refine the class, so the equivalence class is a report on the current design rather than a discovery about the object. Section 8 shows that this has a counterfactual payoff: a refined class carries a narrower ambiguity band.

8 The Set-Valued Lift

This section contains the results original to this paper. Its objective is to determine when the non-identification of Section 7 propagates to the capability object, and when it does not.

The preceding results concern discrimination among hypotheses or latent spaces: questions about which model generated the observed data. The diagnostic's object is different in kind. A reachable viable repertoire is a *set* of trajectories under conditions that were *not realized*. Non-identification of the generating model is neither necessary nor sufficient for non-identification of that set, and no transferred result settles which obtains.

8.1 The Reachability Functional

This subsection introduces the construction that makes the question tractable.

Definition 8.1 (Off-regime reachability functional and repertoire). For a counterfactual condition C' , define

$$f_G^{C'}(\tau) = \sup_{a \in \mathcal{A}(C', h)} K_G(\mathcal{N}_\delta(\tau) \cap \mathcal{V}(C', h) \mid h, a, C'), \quad (1)$$

and

$$\mathcal{P}(G; C') = \left\{ \tau \in \mathcal{V}(C', h) : f_G^{C'}(\tau) \geq \epsilon \right\}. \quad (2)$$

Equation (2) writes the repertoire as a superlevel set. This is the move on which everything below depends: a question about equality of sets becomes a question about how far a scalar functional can move across an equivalence class.

Definition 8.2 (Core and hull).

$$\mathcal{P}^\cap(C') = \bigcap_{G' \in [G]} \mathcal{P}(G'; C'), \quad \mathcal{P}^\cup(C') = \bigcup_{G' \in [G]} \mathcal{P}(G'; C').$$

A trajectory in the core is reachable under every organization consistent with the evidence, so its reachability is entailed. A trajectory outside the hull is reachable under none. Trajectories between them have their status left open. These are the sharp bounds: no procedure using only on-regime evidence narrows them, since every member of the class is consistent with everything observed.

8.2 Maximal Failure Under Off-Regime Surgery

This subsection establishes the first of two brackets on the problem.

Definition 8.3 (Surgery closure). The model class $\mathfrak{G}_d$ is closed under off-regime surgery when for every $G_1, G_2 \in \mathfrak{G}_d$ the splice

$$K_{G_{12}}(\cdot \mid h, a, C) = \begin{cases} K_{G_1}(\cdot \mid h, a, C), & C \in \mathcal{C}_{\text{obs}}, \\ K_{G_2}(\cdot \mid h, a, C), & C \notin \mathcal{C}_{\text{obs}}, \end{cases}$$

again lies in $\mathfrak{G}_d$.

Proposition 8.4 (Collapse of the identified region). *If $\mathfrak{G}_d$ is closed under off-regime surgery and $C' \notin \mathcal{C}_{\text{obs}}$, then for every $G \in \mathfrak{G}_d$,*

$$\mathcal{P}^\cap(C') = \bigcap_{G_2 \in \mathfrak{G}_d} \mathcal{P}(G_2; C'), \quad \mathcal{P}^\cup(C') = \bigcup_{G_2 \in \mathfrak{G}_d} \mathcal{P}(G_2; C').$$

Proof. Fix G and any $G_2 \in \mathfrak{G}_d$, and form the splice G_{12} with $G_1 = G$. It agrees with G at every $C \in \mathcal{C}_{\text{obs}}$, hence induces identical observation laws there under every design, hence lies in $[G]_{\mathcal{Q}_0, \mathcal{C}_{\text{obs}}}$ by Definition 7.1. It agrees with G_2 at C' , so $\mathcal{P}(G_{12}; C') = \mathcal{P}(G_2; C')$. Since G_2 was arbitrary, the equivalence class realizes every repertoire realizable in the whole model class. $\square$

Under surgery closure, on-regime evidence constrains the counterfactual repertoire exactly as much as no evidence at all. This is a genuine failure mode and not a technicality: it obtains whenever the model class permits off-regime behavior to be detached from on-regime behavior.

The result is nonetheless informative. It localizes the entire question in a structural property of the model class, and it shows that any positive result must be purchased by a restriction tying off-regime to on-regime behavior.

8.3 The Ambiguity Band

This subsection establishes the paper's central positive result.

Definition 8.5 (Off-regime deviation).

$$\eta(C') = \sup_{G_1, G_2 \in [G]} \sup_{\tau \in \mathcal{T}} |f_{G_1}^{C'}(\tau) - f_{G_2}^{C'}(\tau)|.$$

Proposition 8.6 (Ambiguity-band sandwich). *For every $G' \in [G]$,*

$$\left\{ \tau \in \mathcal{V}(C', h) : f_G^{C'}(\tau) \geq \epsilon + \eta(C') \right\} \subseteq \mathcal{P}(G'; C') \subseteq \left\{ \tau \in \mathcal{V}(C', h) : f_G^{C'}(\tau) \geq \epsilon - \eta(C') \right\}, \quad (3)$$

and consequently

$$\mathcal{P}^\cup(C') \setminus \mathcal{P}^\cap(C') \subseteq \left\{ \tau \in \mathcal{V}(C', h) : |f_G^{C'}(\tau) - \epsilon| < \eta(C') \right\}. \quad (4)$$

Proof. Write $\eta = \eta(C')$. Suppose $f_G^{C'}(\tau) \geq \epsilon + \eta$. By Definition 8.5, $f_{G'}^{C'}(\tau) \geq f_G^{C'}(\tau) - \eta \geq \epsilon$, so $\tau \in \mathcal{P}(G'; C')$ by (2). This gives the left inclusion. Conversely suppose $\tau \in \mathcal{P}(G'; C')$, so $f_{G'}^{C'}(\tau) \geq \epsilon$. Then $f_G^{C'}(\tau) \geq f_{G'}^{C'}(\tau) - \eta \geq \epsilon - \eta$, giving the right inclusion. Both hold for every $G' \in [G]$ and therefore pass to the intersection and the union of Definition 8.2. A trajectory in $\mathcal{P}^\cup \setminus \mathcal{P}^\cap$ lies in the outer set and outside the inner set, which is (4). $\square$

The interpretation is the paper’s central point. Non-identification of the organization does not spread uncertainty across the trajectory space. It confines uncertainty to a band around the reachability threshold, of width controlled by $\eta(C')$. Outside that band, membership in the counterfactual repertoire is entailed by the evidence despite the organization being unidentified.

The bounds in (3) are computable from a single reference organization together with a bound on η , which is what makes them usable. They are outer approximations to the sharp core and hull rather than the sharp bounds themselves.

8.4 Robust Capability Comparison

This subsection derives the form the diagnostic actually requires.

The diagnostic compares an assessed arrangement with a baseline. Let C^m and C^b be the assessed and baseline conditions, with deviations η_m and η_b .

Corollary 8.7 (Robust disadvantage). *If*

$$\left\{ \tau : f_G^{C^m}(\tau) \geq \epsilon - \eta_m \right\} \subseteq \left\{ \tau : f_G^{C^b}(\tau) \geq \epsilon + \eta_b \right\}, \quad (5)$$

then $\mathcal{P}(G'; C^m) \subseteq \mathcal{P}(G'; C^b)$ *for every* $G' \in [G]$.

Proof. By Proposition 8.6, $\mathcal{P}(G'; C^m)$ is contained in the left set of (5) and $\mathcal{P}(G'; C^b)$ contains the right set. Compose the inclusions. $\square$

A capability disadvantage is identified across the entire equivalence class whenever the gap between the assessed and baseline reachability functionals exceeds the combined ambiguity. This is a signal-to-ambiguity condition. The organization need not be identified; what must be established is that the comparison survives the residual ambiguity.

Corollary 8.7 states a sufficient condition. A disadvantage may be identified through structure this construction does not use, and its failure is therefore not a finding of no disadvantage.

8.5 Counterfactual Distance Bounds the Ambiguity

This subsection establishes the second bracket, quantifying η .

Proposition 8.6 leaves $\eta(C')$ unquantified, and Proposition 8.4 shows it can reach its maximum. Two assumptions control it.

Assumption 8.8 (On-regime channel faithfulness). For $C \in \mathcal{C}_{\text{obs}}$, equality of observation laws under every design in $\mathcal{Q}_0$ implies equality of trajectory kernels: if $\mathcal{L}(Z \mid G_1, C, q) = \mathcal{L}(Z \mid G_2, C, q)$ for all $q \in \mathcal{Q}_0$, then $K_{G_1}(\cdot \mid h, a, C) = K_{G_2}(\cdot \mid h, a, C)$ for all $a \in \mathcal{A}(C, h)$.

Assumption 8.9 (Uniform condition-Lipschitz kernels). There exists $L < \infty$ such that for every $G \in \mathfrak{G}_d$, every h , every action feasible under both conditions, and every $C_1, C_2 \in \mathcal{C}$,

$$\|K_G(\cdot \mid h, a, C_1) - K_G(\cdot \mid h, a, C_2)\|_{\text{TV}} \leq L \rho_{\mathcal{C}}(C_1, C_2).$$

Proposition 8.10 (Extrapolation bound). *Under Assumptions 8.8 and 8.9,*

$$\eta(C') \leq 2L \cdot \text{dist}(C', \mathcal{C}_{\text{obs}}), \quad \text{dist}(C', \mathcal{C}_{\text{obs}}) = \inf_{C \in \mathcal{C}_{\text{obs}}} \rho_{\mathcal{C}}(C', C). \quad (6)$$

Proof. Fix $G_1, G_2 \in [G]$ and $\varsigma > 0$, and choose $C'' \in \mathcal{C}_{\text{obs}}$ with $\rho_C(C', C'') \leq \text{dist}(C', \mathcal{C}_{\text{obs}}) + \varsigma =: r_\varsigma$. Assumption 8.8 gives $K_{G_1}(\cdot \mid h, a, C'') = K_{G_2}(\cdot \mid h, a, C'')$ for every feasible a . Write $A_\tau = \mathcal{N}_\delta(\tau) \cap \mathcal{V}(C', h)$, which is measurable. For any feasible a , inserting the on-regime condition and applying the triangle inequality,

$$\begin{aligned} |K_{G_1}(A_\tau \mid h, a, C') - K_{G_2}(A_\tau \mid h, a, C')| &\leq |K_{G_1}(A_\tau \mid h, a, C') - K_{G_1}(A_\tau \mid h, a, C'')| \\ &\quad + |K_{G_1}(A_\tau \mid h, a, C'') - K_{G_2}(A_\tau \mid h, a, C'')| \\ &\quad + |K_{G_2}(A_\tau \mid h, a, C'') - K_{G_2}(A_\tau \mid h, a, C')| \\ &\leq Lr_\varsigma + 0 + Lr_\varsigma = 2Lr_\varsigma, \end{aligned}$$

where the outer terms use Assumption 8.9 together with the fact that total-variation distance dominates the difference of two measures on any single measurable set. The action set $\mathcal{A}(C', h)$ is common to both organizations, and for real functions u, v with $|u(a) - v(a)| \leq c$ pointwise one has $|\sup_a u - \sup_a v| \leq c$. Applying this to (1) gives $|f_{G_1}^{C'}(\tau) - f_{G_2}^{C'}(\tau)| \leq 2Lr_\varsigma$ for every τ . Taking suprema over τ and over the class, then letting $\varsigma \downarrow 0$, gives (6). $\square$

The ambiguity band grows at most linearly in the distance from the counterfactual condition to the observed regime. Counterfactuals near the observed regime are nearly identified; distant ones are not.

Assumption 8.9 is the structural restriction whose absence Proposition 8.4 exploits, and the two results bracket the problem. Without a tie between off-regime and on-regime behavior the lift fails completely; with a uniform tie it degrades gracefully with distance. The lift therefore admits an informative intermediate answer.

8.6 Consequences for the Surrounding Framework

This subsection records four implications. Each is conditional on the results above and none is an empirical claim.

Baseline admissibility acquires an identification criterion. The companion diagnostic constructs a family of technically feasible arrangements and applies an independent admissibility filter. Proposition 8.10 supplies a reason, internal to identification, to prefer baselines near the observed regime: distant baselines carry proportionally wider ambiguity. This licenses no selection by convenience and does not override the requirement that admissibility be defended independently. It records a cost that distant baselines incur.

Design enrichment has a counterfactual payoff. By Proposition 7.6, enlarging the design family refines the equivalence class. A smaller class has a smaller η , hence a narrower band. Enrichment therefore tightens counterfactual bounds and not merely on-regime discrimination. Choosing which probes to add is an experimental-design problem, and the expected-information criterion supplies the natural objective [9, 3]; computing it for a set-valued target is unresolved.

Existence and magnitude separate formally. Corollary 8.7 supports a comparison without point-identifying either repertoire and without valuing any foreclosed trajectory. This is the formal counterpart of Thesis 5.4.

The unfalsifiability objection acquires a determinate negative. When the assessed functional dominates the baseline functional by more than the combined ambiguity, no member of the class exhibits a disadvantage, and the framework returns a supported negative rather than an unresolved status. This is the formal counterpart of Thesis 5.2.

9 Partial Identification Through Documented Condition Change

This section states a route to a capability claim that bypasses latent identification entirely. Its objective is to supply a result that survives the failure of Section 8, since Proposition 8.4 shows that failure is possible.

Assumption 9.1 (Actor-control monotonicity). Removing feasible actions cannot create actor-controlled reachable branches: if $\mathcal{A}(C^m, h) \subseteq \mathcal{A}(C^b, h)$, then every trajectory reachable and actor-controlled under C^m is reachable and actor-controlled under C^b .

Proposition 9.2 (Documented contraction bound). *Suppose $\mathcal{A}(C^m, h) \subsetneq \mathcal{A}(C^b, h)$ is directly documented, and suppose Assumption 9.1 holds. Then*

$$\mathcal{P}^{\text{aut}}(G; C^m) \subseteq \mathcal{P}^{\text{aut}}(G; C^b)$$

for the actor-controlled sub-repertoires, without point-identifying G and without valuing any foreclosed trajectory.

Proof. Every actor-controlled branch reachable under the contracted arrangement is witnessed by an action surviving the contraction, since the action set under C^m is contained in that under C^b and the kernels are those of the same G . Assumption 9.1 forbids the appearance of new actor-controlled branches under the contracted arrangement. $\square$

The content lies entirely in the two conditions, and both are demanding.

Monotonicity fails under substitution. Constraint provokes adaptation. A foreclosed route can produce a new strategy, a coalition, a change of domain, or a relation that opens branches previously unavailable. This is not an edge case, and a companion project in the same programme treats exactly that substitution as its central mechanism. The tension is real: there, constraint reliably changes the response set, which is what Assumption 9.1 denies here. The assumption is plausible over short horizons and for narrowly specified action sets, and implausible over long horizons where adaptation has time to operate.

Documentation is required, not inference. The contraction must be directly documented. Where feasibility is inferred from behavior, preference adaptation corrupts the measurement [4], and it does so worst under sustained constraint, which is the condition the diagnostic targets. Behavioral inference of feasibility is therefore inadmissible for this purpose. The restriction narrows the applicable domain sharply and is a cost of the result rather than a refinement of it.

10 A Worked Design and Its Limits

This section specifies what an applicable study would require. Its objective is to state requirements, not to demonstrate that they can be met.

The most tractable design uses directly logged access restrictions. Where a controller records access terms, permission grants and revocations, allocation decisions, and prices, the contraction of the feasible action set is observable without inference, which is what Proposition 9.2 requires.

Such a study would proceed in dependency order. Declare the domain, model class, carrier, horizon, viability standard, resolution, and threshold before examining any outcome. Declare the observation-design family and the channel, since the equivalence class is indexed to both. Establish the observed condition regime and the counterfactual condition of interest, together with the metric under which their distance is measured. Report the layer statuses of Table 1 separately. Where Section 8 is used, state the model-class restriction supporting Assumption 8.9 and report sensitivity to plausible alternatives. Where Section 9 is used, exhibit the documentation and defend Assumption 9.1 over the stated horizon.

Three limits apply and none is incidental. The design requires records held by the party whose conduct is under assessment, which is the practical problem Section 6 identifies. The Lipschitz constant is a modeling commitment and no procedure for estimating it is proposed here. And both $f_G^{C'}$ and η are population objects; their estimation from finite samples is untreated, so the results are identification results rather than inference procedures.

11 Open Questions Across Registers

This section collects unresolved questions by the register in which each would be answered.

11.1 Formal Questions

Open Problem 11.1 (Structural status of realistic model classes). Do model classes of interest satisfy Assumption 8.9, or are they closed under off-regime surgery in the sense of Definition 8.3? The positive half of Section 8 is valid in form but vacuous in application if realistic classes fall on the wrong side of this divide.

Open Problem 11.2 (Tightness of the bounds). How loose is (3) relative to the sharp core and hull of Definition 8.2? The bound may be weak when the supremum defining η is attained far from the threshold.

Open Problem 11.3 (Enriched coordinates). The construction bounds the repertoire component alone. Does an analogous lift hold for autonomy, direction, exit, attribution, and the other enriched capability coordinates the companion diagnostic admits?

Open Problem 11.4 (Non-binary discrimination). Proposition 7.2 is stated for binary hypotheses under equal priors. What is the corresponding ceiling for a model class of many candidates, and does the design-family adaptation preserve it?

11.2 Methodological Questions

Open Problem 11.5 (Estimation). Both $f_G^{C'}$ and η are population objects. What estimators exist, what are their finite-sample properties, and does estimation error compound with the ambiguity band or absorb into it?

Open Problem 11.6 (Design selection for a set-valued target). Expected information gain supplies a criterion for choosing probes [9, 3]. What is the corresponding objective when the target is a set and the quantity to be reduced is η ?

11.3 Normative and Institutional Questions

Open Problem 11.7 (Separating the asymmetry from general discovery). Can Thesis 5.3 be sustained as a claim specific to constitutively unobservable objects, or does it reduce to a general argument for discovery obligations?

Open Problem 11.8 (Obligation under permanent unresolvability). What follows for justificatory burden when a layer of Table 1 is not merely unresolved but demonstrably unresolvable under every available design?

11.4 Empirical Questions

No empirical case supports any identification claim in this paper. The priority questions are whether any domain supplies documented condition changes accessible to a complainant; whether a Lipschitz constant can be defended in any applied setting; and whether a determinate negative has ever been produced. A framework that never returns a negative finding in practice would be unfalsifiable regardless of its formal capacity to do so, and Section 12 records this as a defeat condition.

12 Limits and Revisable Research Program

This section consolidates the contribution, states what is transferred and what is new, and fixes the conditions under which the paper should be narrowed or withdrawn.

12.1 Contribution and Its Provenance

The methodological contribution is the six-layer decomposition of Table 1 with its reporting rule and three prohibited promotions.

The negative results of Section 7 are transferred with adaptation from two companion projects and are not claimed as new. What is added there is the adaptation to a design family and the explicit statement of both inferential directions for the generativity setting.

The results of Section 8 are original to this paper: the reachability-functional construction, the core and hull bounds, the surgery-closure collapse, the ambiguity-band sandwich, the robust-comparison corollary, and the extrapolation bound. Their significance is that they answer a question the transferred results leave open, since a set-valued target under counterfactual conditions is not a hypothesis-discrimination problem.

The philosophical contribution is the four theses, of which Theses 5.2 and 5.3 are load-bearing. The political-economic contribution is the structural account of where the usable records sit and the distinction between incidental and constitutive opacity.

The paper establishes no estimator, no general non-identifiability theorem, and no valuation method, and it classifies no observed relation.

12.2 Defeat Conditions

The formal contribution loses support if realistic model classes prove closed under off-regime surgery or admit no uniform Lipschitz constant, since the positive half would then be vacuous in application; if the bounds prove too loose to support any comparison; or if Assumption 9.1 fails generically, removing the bypass route.

The methodological contribution loses support if the layers cannot be evaluated separately even in the weak sense the reporting rule requires, which is that success at one entails success at none.

The normative contribution loses support if Thesis 5.3 cannot be distinguished from a general discovery argument.

The practical contribution loses support if no domain supplies documentation both adequate for Proposition 9.2 and accessible to a complainant, or if no determinate negative is ever produced where the framework is applied.

A narrower paper survives most of these. A rigorous account of what finite traces cannot establish, combined with the layer decomposition and the reporting rule, would still discipline the surrounding programme and would still answer the unfalsifiability objection. The retreat should be stated explicitly rather than performed by weakening claims silently.

12.3 Deferred Questions

The diagnostic itself, its comparison construction, and its benefit-through conditions belong to a companion project. Propagation across relations belongs to another. Legal doctrine, evidentiary burden, remedy, and valuation for adjudication belong to a further project with its own verified sources; this paper supplies only the inferential structure such an account would need, and its use of the lost-chance analogue extends no further than the structural point recorded in Section 5.

The resulting position is conservative. Latent generativity is frequently unidentified, and this paper establishes that trace evidence alone shows neither identity nor independence. It also establishes that this failure need not propagate to the capability object: uncertainty about the organization confines uncertainty about the counterfactual repertoire to a band around the reachability threshold, and a capability comparison survives whenever the gap exceeds that band. Whether real domains satisfy the conditions under which this holds is unresolved, and the paper's most useful outcome may often be an explicit statement of which layer failed and why.

References

- [1] Richard Bellman and Karl Johan Åström. On structural identifiability. *Mathematical Biosciences*, 7(3–4):329–339, 1970.
- [2] Kean Birch. Technoscience rent: Toward a theory of rentiership for technoscientific capitalism. *Science, Technology, & Human Values*, 45(1):3–33, 2020.
- [3] Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995.
- [4] Jon Elster. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press, Cambridge, 1983.
- [5] Robert Givan, Thomas Dean, and Matthew Greig. Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1–2):163–223, 2003.
- [6] Kosuke Imai, Luke Keele, and Dustin Tingley. A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334, 2010.
- [7] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.
- [8] Joseph H. King. Causation, valuation, and chance in personal injury torts involving preexisting conditions and future consequences. *The Yale Law Journal*, 90:1353, May 1981. Cited in legal convention as 90 Yale L.J. 1353 (1981); the final page was not verified and no end page is asserted.
- [9] Dennis V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.

- [10] Charles F. Manski. Nonparametric bounds on treatment effects. *The American Economic Review*, 80(2):319–323, May 1990. Papers and Proceedings.
- [11] Ilya Molchanov and Francesca Molinari. *Random Sets in Econometrics*. Number 60 in Econometric Society Monographs. Cambridge University Press, Cambridge, 2018.
- [12] Ingrid Robeyns. The capability approach: A theoretical survey. *Journal of Human Development*, 6(1):93–117, 2005.
- [13] Amartya Sen. Capability and well-being. In Martha C. Nussbaum and Amartya Sen, editors, *The Quality of Life*, pages 30–53. Oxford University Press, Oxford, 1993.
- [14] Cosma Rohilla Shalizi and Andrew C. Thomas. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, 40(2):211–239, 2011.
- [15] Elie Tamer. Partial identification in econometrics. *Annual Review of Economics*, 2:167–195, 2010.
- [16] Nicholas Vrousalis. Exploitation, vulnerability, and social domination. *Philosophy & Public Affairs*, 41(2):131–157, 2013.
- [17] Matt Zwolinski. Structural exploitation. *Social Philosophy and Policy*, 29(1):154–179, 2012.